Initial Evaluation of USAF Security Forces Distributed Mission Training (SecForDMT)

Dr. Joseph Lloyd Weeks
War Fighter Training Research Division,
United States Air Force Research Laboratory,
6030 South Kent Street, Mesa, AZ, USA 85212-6060
Tel: +1 480 988 6561 Ext 249 Fax: +1 480 988 6285
E-mail: joseph.weeks@williams.af.mil

ABSTRACT

United States Air Force security forces have a central role in force protection. Their missions include military police services, installation security, air base defense, military working dog functions, and combat arms training and maintenance. Surveys indicate that among the hundreds of tasks in the career field, "Directing security forces" is rated highest on training emphasis (Weeks, Garza, Archuleta, and McDonald, 2001). As a result of established needs and technology opportunities, the Air Force Research Laboratory is conducting research and development of a computer-based simulation capability called Security Forces Distributed Mission Training (Weeks and McDonald, 2002). The capability is designed to support training in decisionmaking, leadership, and team coordination. It allows an instructor to start a simulation exercise on trainee computer workstations connected via a local area network. Simulation software supports the interaction of trainees with each other and with computer-generated forces (CGFs) that imitate behavior of enemy, neutral, and friendly troops and civilians. Radio functions allow multi-channel communication among instructors, trainees, and CGFs. A major design objective is to develop a simulation control interface that instructors and trainees can directly use so the costs of an on-site computer technician can be avoided. The purpose of this paper is to describe outcomes from an evaluation of the usability of the simulation control interface and the validity of computer models.

INTRODUCTION

United States Air Force security forces are responsible for military police services, installation security, airbase defense, military working dog functions, and combat arms training and maintenance. Whether security forces discover an improvised explosive device in a car or confront representatives of a non-governmental organization at an entry control point, situations are reported to a command post. Once situations are reported, quick and accurate decisions by security forces leaders are critical to handling the situation properly and fundamental to the protection of personnel and assets.

Command post exercises are routinely conducted to train personnel to respond to diverse situations. Nevertheless, the command and control of security forces continues to be recognized as a high-emphasis training area (Weeks, Garza, Archuleta, and McDonald, 2001). As a result of established needs and technology obstacles, the Air Force Research Laboratory is conducting research and development of a simulation capability for training leadership, decisionmaking, and team coordination (Weeks and McDonald, 2002). An illustration of the current capability is

Paper presented at the RTO HFM Symposium on "Advanced Technologies for Military Training", held Genoa, Italy13 – 15 October 2003, and published in RTO-MP-HFM-101.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **00 APR 2004**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Initial Evaluation of USAF Security Forces Distributed Mission Training (SecForDMT)**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **War Fighter Training Research Division, United States Air Force Research Laboratory, 6030 South Kent Street, Mesa, AZ, USA 85212-6060**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001667, NATO RTO-MP-HFM-101 Advanced Technologies for Military Training (Technologies avancées pour lentraînement militaire)., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**16**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

presented at Figure 1. It consists of standard personal computers connected by a local area network. To reduce acquisition and maintenance costs, visual displays consist of standard computer monitors rather than immersive displays. A trainee station consists of a computer processor, computer monitor, keyboard, mouse, and radio headset and microphone for communications with instructor, other trainees, and computer-generated forces (CGFs). The instructor station consists of one computer for controlling the simulation exercise and recording it for after-action reviews. In Figure 1, trainees are illustrated as a shift leader and subordinate flight leaders but could alternatively be a law enforcement desk sergeant and field officers, a flight leader and subordinate squad leaders, or a defense force commander, operations officer, and field supervisors. The capability is being developed as an expandable, multi-echelon, command and control, training device.

During this initial stage of development, the capability is being tailored to support training for air base defense. During application, an instructor would present on the local area network a terrain model for the area of operations (AO). Trainees would simultaneously view the AO, collaborate in conducting security vulnerability analyses, and develop a security plan. Based on consideration of threats, vulnerabilities, and available resources, trainees would develop the plan by creating and positioning computer models. The simulation capability provides computer models representing sensors, obstacles, fighting positions, vehicles, communications, and semi-automated, CGFs. Computer-generated forces are designed to move, sense, shoot, and communicate. After the plan is completed, trainees would share it with the instructor. The instructor would evaluate the plan and could task CGFs to present security situations to test the plan. After the instructor tasks CGFs to present security situations, the exercise would begin. A situation report from a friendly CGF would be the initiating action for trainee decisionmaking and team coordination. All communications via radio microphones and actions occurring on the visual display would be recorded for after-action reviews. Communications from CGFs to trainees and among trainees would be the focus of student evaluation.

Research objectives include development of a usable control interface, realistic behaviors for CGFs, development of simulation exercises to support learning objectives, and evaluation of system usability, model validity, and training effectiveness. The research and development project includes multiple field evaluations with participation of end users. The development strategy is to collect feedback and apply it to refinement of the capability in an effort to accelerate transition to the field.

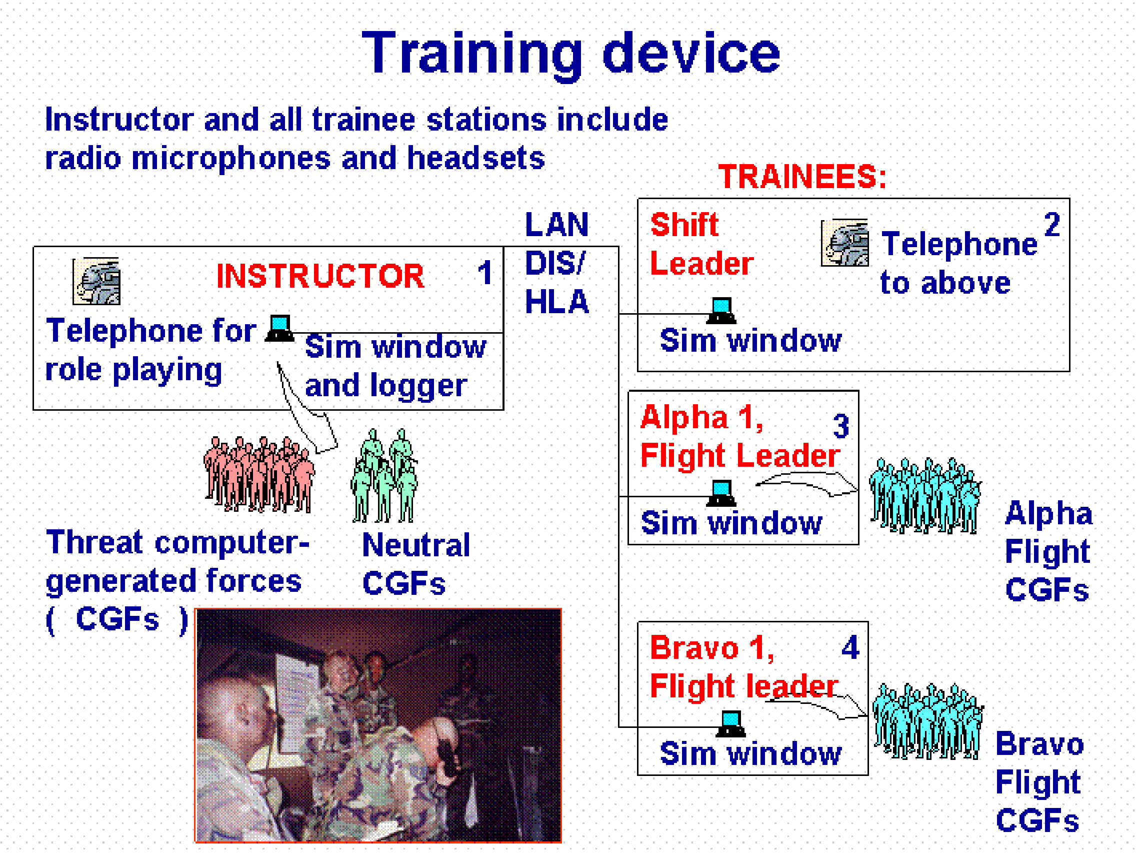


Figure 1. The training device

Usability assessments include measures of time required to train participants to use the capability and instructors' ratings of the usability of the control interface. Instructors' model validity assessments focus on computer models for obstacles, CGFs, and sensors. Evaluations conclude with global ratings of the value of the device for training decisionmaking and team coordination that underlying mission planning and execution of the defense. A preliminary evaluation was conducted and found to be too long in duration (Weeks and McDonald, 2003). As a consequence, the original evaluation procedure was modified. The purpose of this paper is to describe outcomes from the revised procedure for the usability evaluation and future plans for evaluating training effectiveness.

APPROACH

Brewer, Armstrong, and Steinberg (2002) state "usability testing … verifies that a system or product design meets our expectations for usefulness and satisfaction before we move into production (p 403)." They define usability as "the degree to which the design of a device or system may be used effectively and efficiently by a human (p 403)" and point out that the important issue in arriving at a definition is how to measure usability so that measurements can be used to improve the design. Brewer, et al. (2002) outline three general approaches to usability evaluation including surveys (using self-report data collection methods), usability inspections (specialists scrutinize a design according to a systematic approach and judge its acceptability

against certain criteria), and experimental tests (based on quantifying operator performance using controlled data collection techniques).

Usability inspection best describes the approach. The specialists were instructors assigned to different security forces training squadrons. They were first trained to use the simulation control interface; then, they evaluated the interface on selected criteria, identified problems, and recommended improvements. Their evaluation concluded with observation of the performance of computer models and ratings of model validity. These specialists were not experts in human factors engineering, but they did represent a final authority for usability, the end user. Trainees are also end users of the simulation capability. They are expected to control the simulation interface and create computer models for sensors, obstacles, weapons, CGFs, and communications. Trainees participated in the usability evaluation by providing baseline training times for simulation control tasks representative of those they would perform during an exercise. The complete training system consists of several computers linked by a local area network. However, for the usability evaluation only one laptop computer was used. The strategy is to modify the control interface and computer models on the basis of change recommendations before taking the complete training system to the field for evaluation.

To describe the usability inspection approach, Brewer, et al (2002) present an example of a computer interface. It is a single dialogue box. The usability issue is whether to position control buttons on the bottom left or bottom right of the dialogue box. Compared to the dialogue box described by Brewer, et al (2002), the interface evaluated here is huge. It consists of over 450 different controls including menus, tools, dialogue boxes, and intermediate control windows. One recommended approach to usability testing is based on presenting a control interface to end users, not informing them how to use it, observing if they can deduce how to use it, and requesting feedback concerning improvements (Andre, personal communication, 2003). Although this would be a useful approach for the interface described by Brewer, et al. (2002), it was not used here. In addition to the great number of interface controls, most participants had no experience with simulation capabilities like the one evaluated. For them, it was a novel experience; so it was impractical to adopt a discovery learning approach. To minimize the duration of the evaluation period while obtaining meaningful input, it was necessary to familiarize participants with the interface in advance.

The instructor's control interface used for evaluations was delivered with Version 2.0 of the simulation software and is presented at Figure 2. It consisted of a simulation window in which a terrain model is presented; a menu bar consisting of menus, menu options, and menu option labels; a tool bar consisting of tool bar buttons and information windows; a pop-up menu presented in the simulation window and accessed through a right mouse click, a mouse, and a standard computer keyboard. Radio microphone and headset and simulation logger were not included. These interface devices will be tested during evaluation of the complete training system.

One of the greatest obstacles to incorporating such simulations into formal training is support costs. For simulations currently available, on-site technicians are required to design and develop exercises in support of learning objectives and be present during the instructional event to control the exercise, serve as role players, task CGFs, and provide simulation replays to

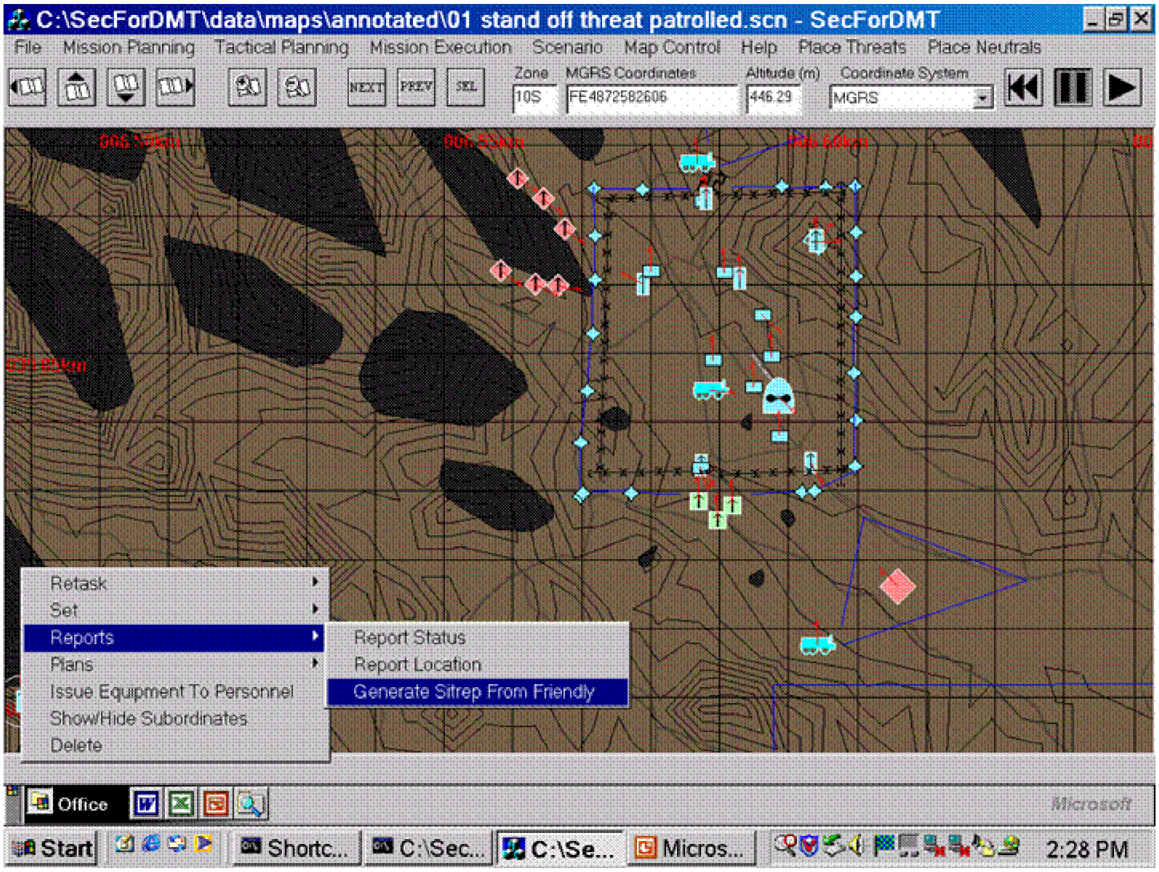


Figure 2. The simulation control interface

support after-action reviews. Support costs are a barrier to fielding simulation technology for formal training; hence, design and development of a simulation control interface that can be directly used by instructors and trainees is a critical technology obstacle. At the beginning of the project, a usability standard was established as the time required to train participants to use the device. The expectation is the shorter the training time; the greater the usability. The intent is to compare the usability standard with observed training times as a guide for system improvement. Training time standards are instructors will be trained to use the device in 2 hours and students will be trained in 30 minutes.

To begin the evaluation, each participant was presented a briefing describing the complete training system and the purpose of the evaluation. After the briefing, they were asked to read and sign a disclosure and consent form and to complete a background questionnaire to obtain information about their time in service, training, and experience with computers. The evaluation was conducted one participant at a time. This approach minimized adverse impact on day-to-day activities of the training group that could have occurred if several participants were tasked to support the evaluation in mass. Each participant was trained to use the control interface and their training time was recorded. It was explained to each participant they were not being evaluated; rather their training time was being measured to estimate training time for the device. Threshold training time was defined as the maximum time that could be allocated for a user to be trained to use the device. After training on all tasks, instructors were asked to estimate maximum training time for instructors and students separately.

The simulation control interface included over 450 control options. Rather than attempting to train participants on all control options, samples of tasks were selected to represent those likely to be performed during a simulation exercise. A task consisted of the use of selected controls to accomplish a purpose. Task selection involved a trade-off between practicality and an exhaustive evaluation. A balance was sought between what participants might regard as an intolerably long evaluation period and evaluation of all control options. The number of tasks was constrained by the length of the evaluation so that the evaluation for each participant did not exceed four hours. Instructors were trained on 18 tasks and students were trained on 13 tasks. Trainee tasks are a subset of instructor tasks and exclude tasks only instructors would perform like creating threat and neutral CGFs. Instructor and trainee tasks are non-random samples from the population of tasks. Although the main purpose was to evaluate usability, the evaluation was the first opportunity for instructors to observe the training device; so a secondary goal was to make the experience meaningful. Rather than randomly sampling and presenting tasks, they were carefully selected and sequenced for meaningfulness. Tasks for both instructors and trainees represent three categories consisting of control of the simulation window, placing resources, and tasking CGFs. If more tasks had been selected, total training time would have been greater. However, the more tasks selected; the greater the length of the evaluation period, and the less likely it would be to obtain participants' willing cooperation and meaningful input. Instructor and trainee tasks and control options used for the evaluation are described in the Appendix.

Measures of training time were obtained separately for each participant. The participant was told how to perform each task, she or he was showed how to do it, and asked to independently perform the task with assistance. They were asked to indicate when they had learned to use controls for the task. When the participant stated she or he had learned to perform the task, training time was declared complete. Immediately after training for each task, the participant was asked to perform the task independently. It was noted whether it was performed with or without assistance. It was assumed that if the participant satisfactorily performed the task without assistance, they had learned to use the controls. If the participant asked for assistance, recorded training time for that task for that participant was doubled. Task training time was cumulated over all tasks to obtain an estimate of training time for each participant. Training time was averaged over trainees and instructors separately to estimate total training time for each group.

Trainees did not evaluate the device. Evaluations were conducted only by instructors. After training for each task, the instructor rated the controls used to perform the task on clarity, effectiveness, efficiency, and simplicity; always in that order. Clarity was defined as the degree to which interface controls were clear and understandable. Effectiveness was defined as the degree to which interface controls allowed the task to be performed. This factor provided an opportunity for instructors to recommend needed functionality. Efficiency was defined as the degree to which the controls used in performing the task allowed quick performance. Simplicity was defined as the degree to which the logic of using the controls was complex or easy to understand. Extremes of the rating scale for simplicity were anchored with verbal anchors, "Extremely High Simplicity" and "Extremely High Complexity". Rating scales for clarity, effectiveness, and efficiency were like the one described in Figure 3 except the applicable factor was inserted. The process of obtaining ratings guided the instructor to think about specific criteria for usability and provided indicators of order relationships among tasks for each factor. After rating the control interface for a task on a factor, instructors were asked to identify problems and recommend improvements.

0 = "Do not know" 1 = "No <u>clarity</u> " 2 = "Extremely Low <u>clarity</u>" 3 = "Very Low <u>clarity</u>" 4 = "Below Average <u>clarity</u>" 5 = "Average <u>clarity</u>" 6 = "Above Average <u>clarity</u>" 7 = "High <u>clarity</u>" 8 = "Very High <u>clarity</u>" 9 = "Extremely High <u>clarity</u>".

Figure 3. Example of one of four usability rating scales

When task training and usability ratings were completed, each instructor evaluated the validity of computer models for obstacles, CGFs, and sensors. They observed five simulation runs that showed the performance of selected computer models and rated the validity of what they observed. The validity rating scale was like the one described in Figure 3 except the word, "Validity", was inserted for the underlined factor. In addition, a sensor expert and a weapons expert reviewed specifications used to develop computer models for sensors and weapons and corrected specifications as appropriate. After observing simulation runs, instructors were asked if they believed the device would support learning objectives for mission planning and execution of the defense and whether or not they believed the simulation capability would add value to training.

RESULTS

A total of 13 security forces instructors participated in the evaluation (3 instructors from the 342nd Training Squadron, 7 instructors from the 343rd Training Squadron, and 3 instructors from the 96th Security Forces Ground Combat Training Squadron). Each instructor dedicated approximately 4 hours to the evaluation including breaks. Participants also included 10 trainees from the 343 Training Squadron who had recently graduated from security forces initial-skills training and were awaiting assignments.

On average, the instructors were 31 years of age. There were 3 captains, 1 senior master sergeant, 1 master sergeant, 4 technical sergeants, and 4 staff sergeants. All enlisted personnel serving as instructors possessed the journeyman skill level or higher. They had an average of 11 years and 6 months of service and 2 years and 2 months in their current position. They indicated they spend an average of 35 hours per week using computers and in the preceding year played computer games an average of 3 times. On average, trainees were 19 years of age. Five trainees possessed the rank of airman first class and 5 possessed the rank of airman basic. All trainees possessed the apprentice skill level. They had an average of 6 months of service. They indicated they spend an average of 4 hours per week using the computer and in the preceding year played computer games an average of 3 times.

Figure 4. Average task training time for 13 instructors
(Summation over 18 tasks = 1 hr 17 min)

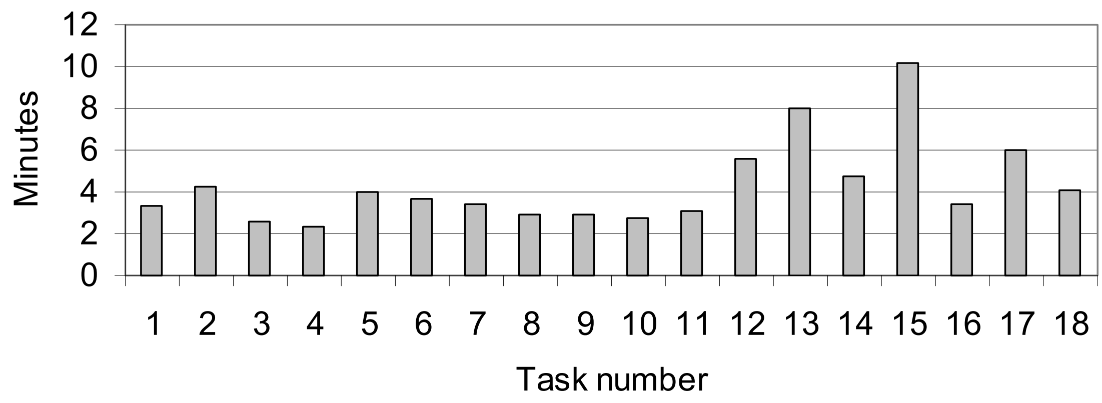


Figure 5. Average task training time for 10 trainees
(Summation over 13 tasks = 56 min 2 s)

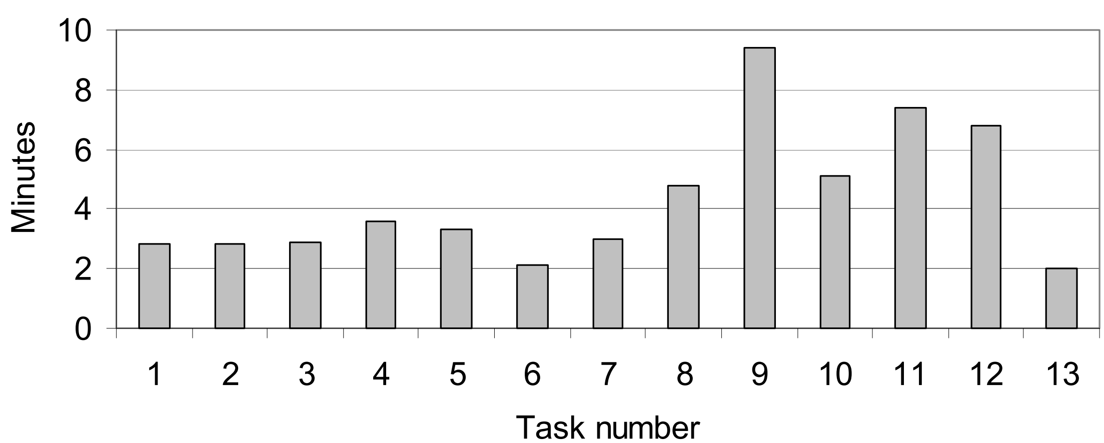
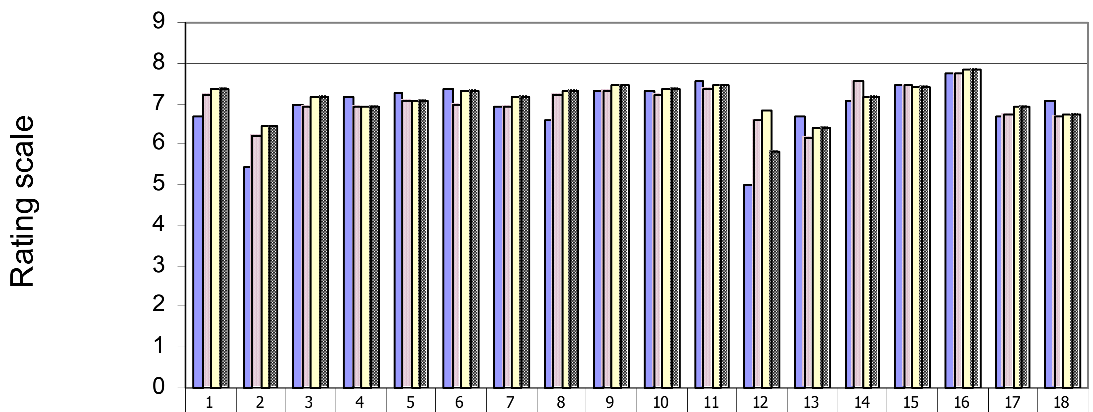


Figure 4 presents the observed task training time for 18 tasks averaged over instructors. Most tasks required 5 minutes or less training time. The instructor training time usability standard was 2 hours. Training time for instructors summed over the 18 task sample is 1 hour and 17 minutes. Average training time for the 18 task sample fell below the pre-established standard. After task training, each instructor was asked to estimate the maximum amount of time that could be allocated for instructors to be trained to use the device. The minimum was 1 hour, the maximum was 8 hours, and the median was 4 hours.

Figure 5 presents the observed training time for 13 tasks averaged over trainees. Most of the tasks required 5 minutes or less training time. For trainees, the training time usability standard is 30 minutes. Training time for students summed over the 13 task sample is 56 minutes. Average student training time for the 13 task sample exceeded the pre-established standard. Each instructor was asked to estimate the maximum amount of time that could be allocated for students to be trained to use the device. The minimum was 1 hour, the maximum was 8 hours, and the median was 4 hours.

For instructors, the procedure was to learn to use interface controls then rate the control interface on clarity, effectiveness, efficiency, and simplicity. Figure 6 presents usability ratings for each factor. Instructors were told that the collection of interface controls used to perform the task were the target for each rating. For each usability factor, a rating value of 5 represents an "Average" rating. Aggregate ratings for all tasks and factors were rated average or above. These

Figure 6. Average interface usability



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Clarity	6.7	5.5	7.0	7.2	7.3	7.4	6.9	6.6	7.3	7.3	7.5	5.0	6.7	7.1	7.5	7.8	6.7	7.1
Effectiveness	7.2	6.2	6.9	6.9	7.1	7.0	6.9	7.2	7.3	7.2	7.4	6.6	6.2	7.5	7.5	7.8	6.8	6.7
Efficiency	7.4	6.5	7.2	6.9	7.1	7.3	7.2	7.3	7.5	7.4	7.5	6.8	6.4	7.2	7.4	7.8	6.9	6.8
Simplicity	7.4	6.5	7.2	6.9	7.1	7.3	7.2	7.3	7.5	7.4	7.5	5.8	6.4	7.2	7.4	7.8	6.9	6.8

Task

results indicate instructors believed the control interface was easy to understand and effective. It allowed them to accomplish tasks quickly and the logic of operations required for using the controls was easy to understand. Ratings for Task 16 deserve special attention. This task presented the capability for instructors to send situation reports from friendly CGFs to serve as a stimulus for trainee decisionmaking and team coordination. Control options allowed the instructor to open a dialogue box with an editable window in which a situation report could be typed/stored and immediately sent from a selected friendly CGF. The process of sending the situation report revealed the sound dimension of the simulation capability, highlighted audible communications among trainees and CGFs, and illustrated support for training decisionmaking and team coordination. Task 16 was rated highest on usability. Even though instructors provided favorable usability ratings for the interface, they identified several problems with the interface and made important recommendations for improvement.

Figure 7. Frequency of problems/recommendations summed over 13 instructors (Total = 204)

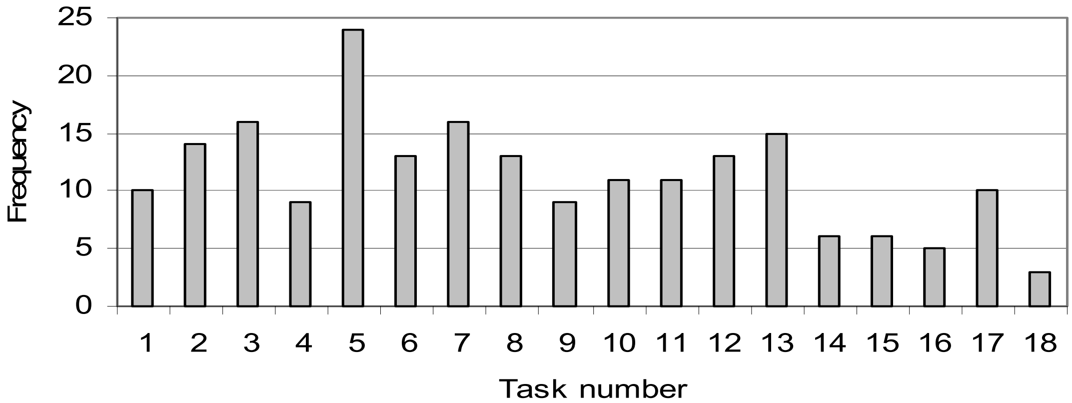


Figure 7 presents the frequency of interface problems/recommendations summed over instructors for each task. They provided a total of 204 change recommendations. Change recommendations were immediately consolidated, prioritized, and shared with development engineers for use as a guide for system development. Examples of change recommendations are presented below:

- Move as many options as possible from tool bar menus to right-click menus.
- Provide a reference point for determining the level of zoom.
- Make it easier to select the end of a range and bearing line.
- Reduce the number of steps required to display terrain contours.
- Reduce the number of steps required to create sensors, CGFs, etc.
- Collapse aircraft menu options into aircraft priority level options.
- Provide an information window to allow instructors to quickly determine the number of weapons, vehicles, radios, and sensors used in the security plan.
- Provide short-cut index for instructor to review all situation reports at once without having to push buttons to display each stored situation report.

After training and usability ratings, instructors observed simulation runs that presented the performance of computer models for obstacles, CGFs, and sensors. Simulation runs included (1) <u>Obstacle delay</u>: A lead, dismounted CGF was tasked to move to a waypoint beyond a single strand of concertina wire and two CGFs were tasked to follow; (2) <u>Obstacle avoidance</u>: A vehicle was positioned in front of a fence and was tasked to move to a waypoint beyond the fence; (3) <u>Fire at will</u>: A dismounted, friendly CGF armed with a M16 was tasked to wait indefinitely with rules of engagement set to "Fire at will'. Rules of engagement for a dismounted, threat CGF armed with an AK47 were set to "Fire at will" and it was tasked to move to a waypoint immediately behind the friendly CGF; (4) <u>Fire if fired upon</u>: A dismounted, friendly CGF armed with a M16 was tasked to wait indefinitely with rules of engagement set to "Fire if fired upon". Rules of engagement for a dismounted, threat CGF armed with an AK47 were set to "Fire at will" and it was tasked to move to a waypoint immediately behind the friendly CGF; (5) <u>Sensors</u>: A dismounted CGF moved on a route past a sensor beam for an active infrared sensor, a trip wire, a

passive infrared sensor, a seismic sensor, and a magnetic sensor. In addition, a truck and tank moved on separate routes that led them within 50 meters, 25 meters, and 1 meter of seismic and magnetic sensors.

Figure 8. Computer Model Validity

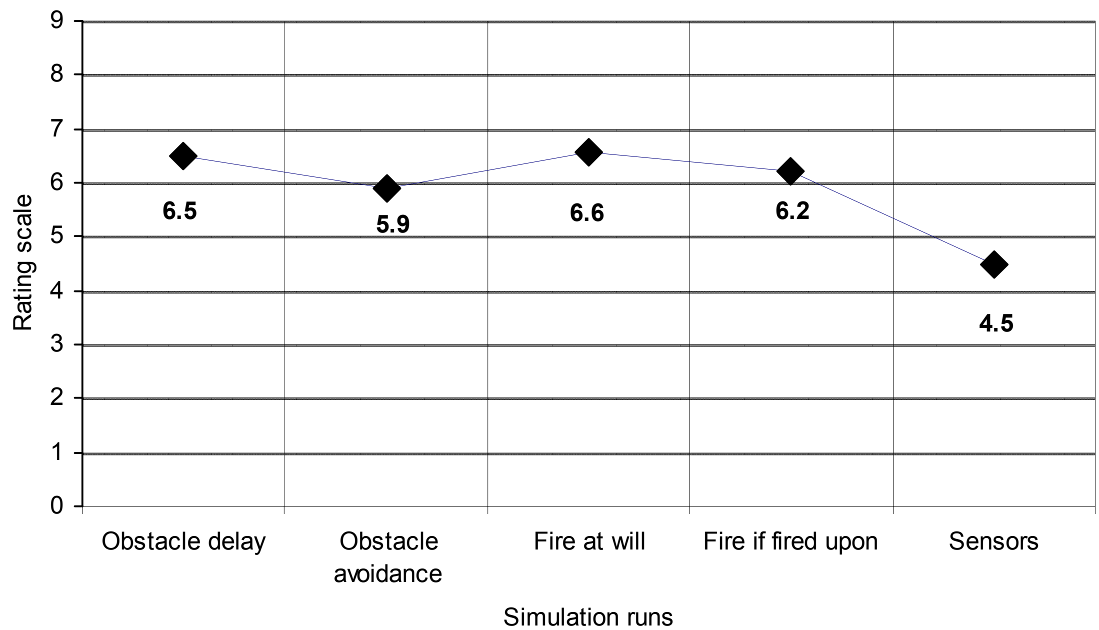


Figure 8 presents average computer model validity ratings for each simulation run. A rating value of 5 represents an "Average validity" rating. Validity ratings for all simulation runs were above average except for sensors. The performance of sensor models was questioned because seismic and magnetic sensors did not detect the presence of vehicles and did not detect the presence of a dismounted CGF at expected distances.

Although there were no weapons validity ratings obtained, the "Fire at will" and "Fire if fired upon" simulation runs did showcase weapons performance. Several instructors commented on weapons effects. For the "Fire at will" simulation run, a friendly CGF armed with a M16 killed the threat CGF armed with an AK47 at approximately 500m. More than one instructor stated that a kill at 500 meters by an average shooter was unrealistic. Under the "Fire if fired upon" condition, a friendly CGF armed with a M16, standing in the open, not in defilade, was hit by a threat CGF armed with an AK47 at 100 meters. The friendly CGF returned fire and killed the threat CGF. The weapons expert stated the effect of the AK47 on the friendly CGF was unrealistically minor and that, in reality, the AK47 would have a more damaging effect. Discussions with software engineers revealed the probability of kill for friendly CGFs was based on the assumption they would be wearing protective Kevlar vests. Revised kill probabilities were obtained and passed to development engineers for system revisions.

After task training and the process of providing usability and validity ratings, each instructor was asked to provide global evaluations. All 13 instructors agreed the device would add value to training and would support learning objectives for security planning and execution of

the defense. The officer responsible for managing the security forces officers' command course expressed the desire to have the simulation capability available for the course by the next class.

DISCUSSION

Although training times provide useful baseline estimates of usability, they have limitations. The evaluation did not include the simulation logger or the radio microphone or headset. If tasks for these devices had been included, training time would have been greater. Training time for these devices will be estimated when the complete training system is evaluated. Because it was necessary for reasons of practicality to limit the number of tasks evaluated, training times are underestimates. Only a representative subset of 18 tasks was evaluated for instructors and only 13 tasks were evaluated for trainees. Observed training times are underestimates of total training time for all 450 interface control options. However, knowledge of all control options is not necessary for practical use. How many users of Microsoft Word know how to apply all control options? For instructors, total observed instructor training time for the 18-task sample is an overestimate. During the training period for each task, instructors often discussed problems and recommended needed capabilities. The result was that total observed training time includes time for training and non-training interactions. It would be desirable to obtain a more accurate estimate of training time that excludes time for non-training interactions, includes tasks for use of the radio headset, and simulation logger, and a greater number of representative tasks while maintaining a reasonable time for the evaluation period. Multiple evaluation periods may be necessary if acceptable to participants.

Apart from limitations in observed training time, it is important to understand the potential represented by this capability. If observed training time for instructors were doubled, it would still be less than the training time required for other simulation capabilities available today. Consider the Joint Combined and Tactical Simulation (JCATS). On a separate occasion, the author observed security forces personnel being trained on JCATS to support installation security evaluations. Observation of "user training" for JCATS indicated 7 security forces personnel varying in rank from staff sergeant to master sergeant satisfactorily learned to task CGFs to move, dismounted and mounted, and shoot, direct and indirect-fire weapons, in 1 and ½ eight-hour, training days. JCATs controllers indicated they had learned to operate JCATS over a period of months. If the training time for JCATS is the comparison point, SecForDMT offers a significant advantage by avoiding lengthy training times for users.

Computer model validity assessments indicated instructors believed computer models for obstacles and CGFs were above average in validity. The simulation run that illustrated obstacle performance showed dismounted CGFs being automatically delayed while moving past a strand of concertina wire and a vehicle automatically avoiding a fence strand. The simulation run that illustrated CGF behavior showed a friendly dismounted CGF complying with rules of engagement for "Fire at will" and "Fire when fired upon". The aggregate rating for sensors was below average. This reflects the fact that seismic sensors did not alarm at the expected detection distance and magnetic sensors did not alarm for vehicles.

The usability evaluation provided valuable information for spiral-development; but, no information describing training effectiveness. Estimates of training effectiveness would be required to quantify benefits for conducting cost-benefits analyses which, in turn, would inform acquisition decisions. Estimates of training effectiveness could be obtained by conducting an experiment using controlled data-collection techniques. An initial objective would be to determine if decisionmaking performance at the individual level improves over successive

simulation exercises. To reduce resource requirements, a repeated-measures, pre-test/post-test, experimental design would be used to attain the greatest statistical power for the minimum number of participants. At least, 20 individuals would be required to participate. The value of the training device can not be separated from the instructor who applies it and the simulation training syllabus. So, it would be necessary to enlist instructors' participation to develop simulation exercises, define performance standards, deliver exercises, and evaluate trainee performance. Resource requirements would be significant. Procedures for conducting such an experiment are currently being formulated and discussed with representatives of training organizations.

CONCLUSIONS

The capability holds promise for being a highly usable training device. Observed training time satisfied the standard for instructors but failed to satisfy the standard for trainees. Although trainees' observed time of 56 minutes exceeded the goal of 30 minutes, the maximum time judged available for training students to use the device (median of 4 hours) suggests the capability could be assimilated into training. Instructors rated usability average or above average for all tasks and rated the validity of computer models above average for all computer models except sensors. Instructors unanimously agreed the device would add value to training and that it would support learning objectives for security planning and execution of the defense. Most importantly, the usability evaluation resulted in over 200 change recommendations. While the system is being improved in accordance with these recommendations, a plan is being formulated to take the complete training system to the field for the purpose of identifying relevant learning objectives, developing simulation exercises, and identifying performance standards and metrics. This information will inform design of an experiment to quantify training effectiveness.

REFERENCES

Armstrong, S.D. and Steinberg, R.K. (2002) *Usability Testing.* In Charlton, S.G., O'Brien, T.G. (Eds.) H*andbook of Human Factors Testing and Evaluation.* Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, Chapter 18, p 403-432.

Weeks, J. L., Garza, J. S., Archuleta, M. A. and McDonald, L. B. (2001) USAF Security Forces Training Needs. *Proceedings of the 2001 Interservice / Industry Training, Simulation and Education Conference*, Orlando Florida.

Weeks, J. L. and McDonald, L. B. (2002). USAF Security Forces Distributed Mission Training: Evolution of the design concept. In, *Proceedings of Spring 2002 Simulation Interoperability Workhop*, Orlando, FL.

Weeks, J. L. and McDonald, L. B. (2003). Preliminary evaluation of the control interface for a training simulation system. *Proceedings fo the North Atlantic Treaty Organization, Research and Technology Agency, Systems Concepts and Integration Panel, Symposium on Critical Design Issues for the Human-Machine Interface,* Prague, Czech Republic, 19-21 May, 2003.

APPENDIX

INSTRUCTOR TASK LIST (K_i=18)

Task module: Control simulation window

Task 1. Zoom in 4 levels, zoom out 2 levels, pan right, and use Sel/Next/Prev buttons to center map.
Task 2. Determine distance represented by side of a map grid square by interpreting map coordinates.
Task 3. Create, select, move, and orient a range and bearing line and determine distance represented by side of map grid square and delete line.
Task 4. Turn on contour line display and determine elevation with cursor and Altitude window.

Task module: Resource placements

Task 5. Create, select, move, and delete a facility and an aircraft.
Task 6. Create, select, move, and delete an obstacle.
Task 7. Create, select, move, orient, and delete a sensor.
Task 8. Create, select, move, orient, and delete a friendly computer-generated force (CGF) armed with an M9 and a friendly CGF armed with an M2.
Task 9. Create, select, move, orient, and delete a threat CGF armed with an AK47 and a threat vehicle (BMP-2 AFV).
Task 10. Create, select, move, orient, and delete an armed neutral CGF and an unarmed neutral CGF.
Task 11. Create, select, move, orient, and delete a defensive fighting position.
Task 12. Place friendly CGF in defensive fighting position, turn on primary field of fire and effective range and set primary field of fire.
Task 13. Create a terrain box, create an observation point, determine intervisibility from the observation point, delete the terrain box and observation point.
Task 14. Withdraw and assign telephones to facilities and set loop designations.
Task 15. Withdraw and assign radios to entities, set call signs, and receive and transmit frequencies.
Task 16. Create, save, and send a situation report.

Task Module C: Control Forces

Task 17. Create a route and write a task plan for a HMMWV to patrol a route at maximum speed.
Task 18. Retask.

TRAINEE TASK LIST (K_t=13)

Task module: Control simulation window
Task 1. Use Sel/Next/Prev buttons to center map on entities, zoom in 4 levels, zoom out 2 levels, and pan.
Task 2. Create, select, move, and orient a range and bearing line and determine distance represented by side of map grid square and delete line.

Task 3. Turn on contour line display and determine elevation with cursor and Altitude window.

Task module: Resource placements

Task 4. Create, select, move, and delete an obstacle.
Task 5. Create, select, move, orient, and delete a sensor.
Task 6. Create, select, move, orient, and delete a friendly computer-generated force (CGF).
Task 7. Create, select, move, orient, and delete a primary fighting position.
Task 8. Place friendly CGF in primary fighting position, turn on primary field of fire and effective range, and set primary field of fire.
Task 9. Create a terrain box over the primary fighting position, create an observation point, determine intervisibility from the observation point, delete he terrain box and observation point.
Task 10. Withdraw and assign field telephones and set loop designations.
Task 11. Withdraw/assign radios, set call signs, and channels, enable audible radio communications, create reference point, and send a location report.

Task Module: Control Forces

Task 12. Create a route and write a task plan for a HMMWV to patrol a route at maximum speed.
Task 13. Use the Retask menu option to retask an entity to go to a particular point on the terrain map.